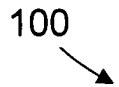
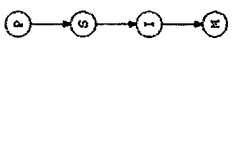
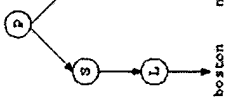



```
⟨!ELEMENT purchases   (purchase*)⟩
⟨!ELEMENT purchase    (seller, buyer)⟩
⟨!ATTRIST seller      ID ID location CDATA name CDATA⟩
⟨!ELEMENT seller      (item*)⟩
⟨!ATTRIST buyer       ID ID location CDATA name CDATA⟩
⟨!ELEMENT item        (item*)⟩
⟨!ATTRIST item        name CDATA manufacturer CDATA⟩
```

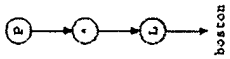
Figure 1
(Prior Art)



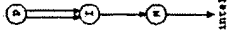
(a) Q_1: find all manufacturers that supply items



(b) Q_2: find orders with Boston sellers and NY buyers



(c) Q_3: find orders with a Boston seller or buyer



(d) Q_4: find orders that contain Intel products (items or subitems)

Figure 2 (Prior Art)

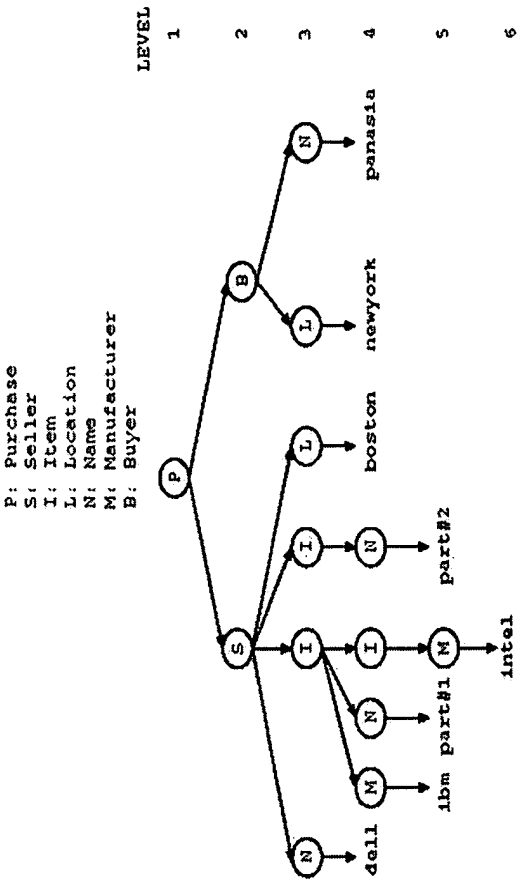

P: Purchase
S: Seller
I: Item
L: Location
N: Name
M: Manufacturer
B: Buyer
LEVEL
1
2
3
4
5
6
P
S
B
I
L
N
M
panasia
newyork
boston
part#2
part#1
intel
ibm
dell

Figure 3

$$\mathcal{D} =$$
$$\underline{(P,\epsilon)}, \underline{(S,P)}, (N,PS), (v_1,PSN), (I,PS), (M,PSI),$$
$$(v_2,PSIM), (N,PSI), (v_3,PSIN), (I,PSI), (M,PSII),$$
$$(v_4,PSIIM), (I,PS), (N,PSI), (v_5,PSIN), \underline{(L,PS)},$$
$$\underline{(v_6,PSL)}, \underline{(B,P)}, \underline{(L,PB)}, \underline{(v_7,PBL)}, (N,PB), (v_8,PBN)$$

Figure 4

Doc_1 : $(P, \epsilon)(S, P)(N, PS)(v_1, PSN)(L, PS)(v_2, PSL)$

Doc_2 : $(P, \epsilon)(B, P)(L, PB)(v_2, PBL)$

Q_1 : $(P, \epsilon)(B, P)(L, PB)(v_2, PBL)$

Q_2 : $(P, \epsilon)(L, P*)(v_2, P*L)$

(a) XML docs and queries in structure-encoded sequences

(b) A tree structure for Doc_1 and Doc_2

Figure 5

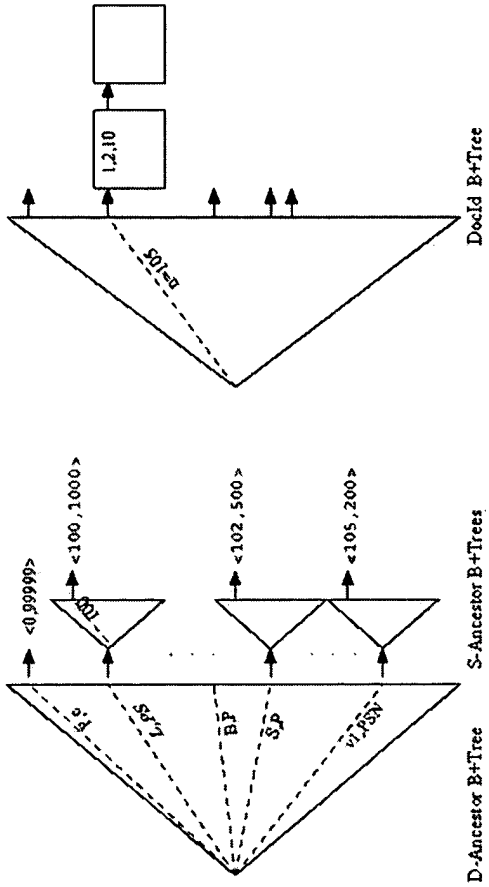

1,2,10
DocId B+Tree
<100,1000>
<102,500>
<105,200>
<0,99999>
S-Ancestor B+Trees
D-Ancestor B+Tree

Figure 6

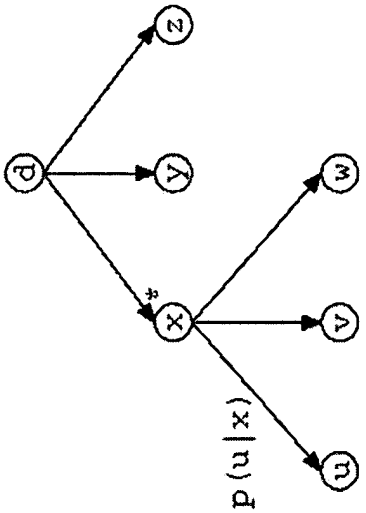


Figure 7

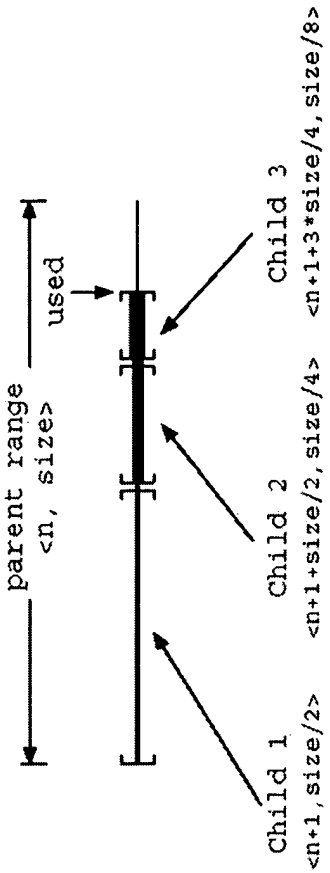
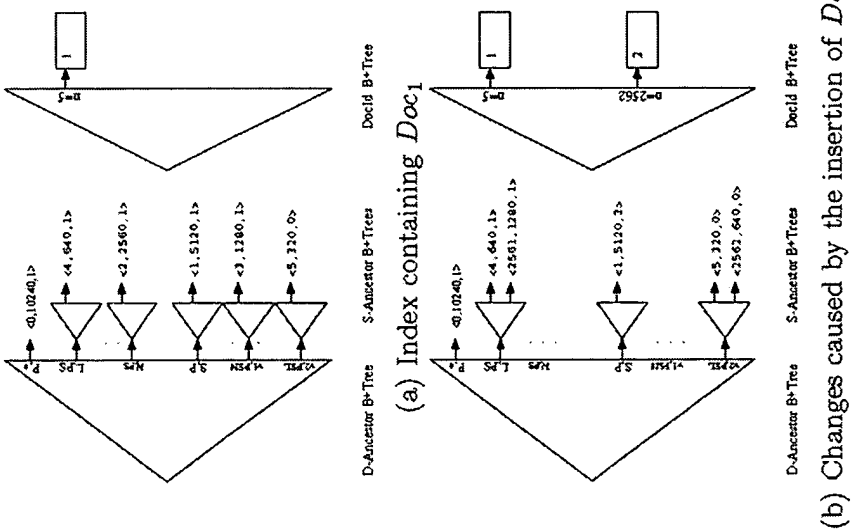

parent range
<n, size>
used
Child 1
<n+1, size/2>
Child 2
<n+1+size/2, size/4>
Child 3
<n+1+3*size/4, size/8>

Figure 8



(a) Index containing Doc_1

(b) Changes caused by the insertion of Doc_2

Figure 9